2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)
Amity University, Noida, India. Mar 14-15, 2024

2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) | 979-8-3503-5035-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICRITO61523.2024.10522438

Devanagari Ancient Manuscript Recognition Using AlexNet

Aditi Moudgil
Chitkara University Institute of Engineering and Technology
Chitkara University
Punjab, India.
aditi.modgil@chitkara.edu.in

Saravjeet Singh
Chitkara University Institute of Engineering and Technology
Chitkara University
Punjab, India.
saravjeet.2009@gmail.com

Bhavna Sareen
Chitkara University Institute of Engineering and Technology
Chitkara University
Punjab, India.
bhavna.sareen@chitkara.edu.in

Shivani Wadhwa
Chitkara University Institute of Engineering and Technology
Chitkara University
Punjab, India.
shivani.wadhwa@chitkara.edu.in

***Abstract*—Ancient manuscripts are one of the major source of obtaining valuable information. Reading these manuscripts is a tedious task so there came a requirement of digitization of these images. Many such OCRs exist which can decode the manuscripts and convert them into a human readable form. One such OCR is developed in this paper which was able to read regional language- Devanagari. For this classification AlexNet transfer learning model is used. In this paper Devanagari manuscript data set is used to classify the Devanagari characters. A total of 7356 characters were included in the dataset which was a combination of Devanagari basic characters and some modifiers and conjuncts. The experiments were performed after dividing the dataset into 90-10 (P1), 80-20 (P2) and 70-30(P3) train-test data. A maximum accuracy of 95.4% was obtained when the model was made to run on P1 partitioning strategy on 30 epochs. The average accuracy of 93.3% obtained for AlexNet model. Also, this paper compares the performance of other CNN based OCRs.**

***Index Terms*—Devanagari, AlexNet, Old Handwritting, Handwriting Recognition, Transfer Learning, Benchmarking**

I. Introduction

The digitization of recording has brought about significant advancements, making text recognition and conversion vital in today's world. In the past, record-keeping relied on handwritten documents known as manuscripts, representing the original versions crafted by authors. Before the era of digital computing, various important materials such as land records, personal histories, holy books, bills, and reading materials were all composed using handwritten text. Preserving these ancient handwritten texts and documents holds immense value for future generations [1]. As technology continues to advance, there is an increasing need to convert these documents into digital formats. Extensive research and development have been dedicated to this field for many years. While there are numerous tools and techniques available for converting handwritten text to digital forms in popular and international languages, there has been limited progress in this area for regional languages [2].

India, known for its rich cultural heritage and vast diversity, encompasses a multitude of languages and scripts for communication and documentation. In this research paper, we focused on the recognition and classification of Devanagari manuscripts. Devanagari scripts serve as the base script which is used to write other languages like Hindi, Nepali, Sanskrit and many more. Devanagari- a combination of two words "Deva" and "Nagari" means "City of God". The scripts evolved from "City of God" are likely to be traditional and old scripts which contain valuable information. Devanagari is a script belonging to the Brahmi script family [3]. Details classification scripts are shown in Fig 1. Devanagari script

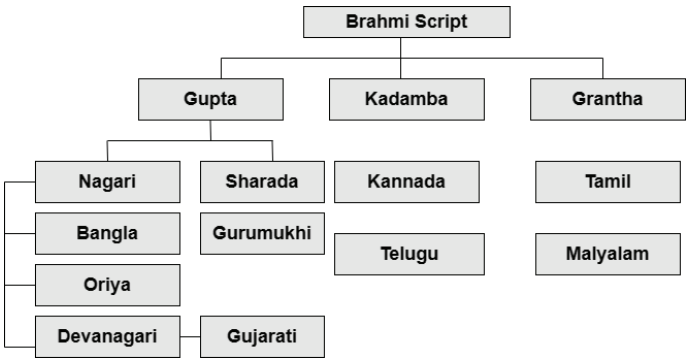

Fig. 1. Evolution of Indian scripts from Brahmi script

follows a left-to-right writing direction, where each word is marked by a header line positioned at the top. In this script, each consonant inherently carries a vowel, and additional vowel diacritics are utilized to modify the representation of consonants and convey specific phonetic qualities. Diacritical marks in Devanagari can appear below, above, to the left, or to the right of consonant letters. The Devanagari script comprises 11 vowels and 33 consonants, collectively referred to as basic characters [4].Basic characters of Devanagari are shown in Fig 2. Devanagari characters exhibit a wide range of

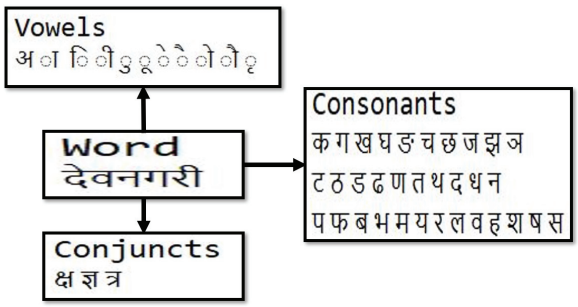

Fig. 2. Elements of Devanagari Script

variations and complexities, resulting in diverse forms of Devanagari manuscripts. Optical Character Recognition (OCR) is a technique used to interpret handwritten and printed text. Although OCR solutions for Devanagari manuscripts do exist, their accuracy and efficiency, particularly for old documents, are often inadequate [5]. In Fig 3 the quality of text being low, due to aging the characters becomes invisible. Character recognition becomes difficult in this case [6].

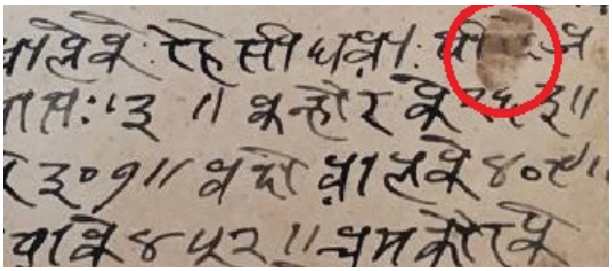

Fig. 3. Sample document for Devanagari Manuscript

II. Literature Study

Lot many research papers have been published in the field of OCR. Initially the OCRs worked only for the recognition of some specific languages and fonts but with the advancement in technology newer OCRs have been developed which perform with upto 99% accuracy for printed texts. But still the need lacks for OCR capable of reading handwritten texts, especially for Devanagari scripts.

M. Buchler emphasized Coptic texts ancient Egyptian scripts. Two different fonts of Coptic texts were experimented upon (Bohairic and Sahidic). The character sets that have been experimented with were taken from several published manuscripts and the problems faced were the different frequency of characters, punctuations, page breaks and line indications [7]. The Coptic font pages were divided into testing and training data such that each page was considered as a part of training data. For pre-processing Scan Tailor method was used to eliminate the borderlines and page cutting. The output of pre-processing was tested by the coptologist and a transcript was generated. The transcript so generated was proofread and the necessary corrections were made. 30,000 training steps were carried out to train the system. Dhir et al. in 2012 presented a new method for the recognition of Devanagri documents. In this method for line segmentation, an average line-height of 30px was considered due to the variation in the handwriting of different people [8]. Here line segmentation was based on the horizontal projection method. The Word segmentation process was considered somewhat easier as the minimum 3 px distance was considered in between two words. In zone segmentation actual header line was compared to the expected header line and subsequently, the header line was straightened and the three segments were separated upper modifier, header line, and other characters.

Singla et al. developed an OCR which was able to recognize Odia characters. CNN based techniques were used for feature extraction and classification [9], [10]. Various methodologies have been explored and advanced within the research community, such as Gaussian filter techniques, zoning, bagging, structural approaches, and statistical feature extraction techniques, as documented in previous studies [11]. Additionally, a variety of classifiers, including CNN, Support Vector Machine (SVM), fuzzy set among others, have been employed. This section presents relevant studies that align with and contribute to these diverse techniques.

Different authors have proposed techniques for the text segmentation and out of these techniques is line segmentation that is used to segment the page into narrower segments called line segmentation [12]. Theeramunkong et al. worked on a method named as island based projection that worked multi-dimensional and extracted some global features from handwritten scripts. Two statistical approached named as n-gram and hidden Markov model were used. The authors used Multi-layer Perceptron (MLP) which performs better while it is implemented on good quality features [13]. The researchers do not need to explicitly define the feature in deep learning but the developed model automatically extracts the best features followed by the input classification. Convolutional neural network, a deep model works with very a smaller number of parameters. Each layer has some parameters which are varied to get accurate results. Also, the number of layers can increased or decreased to get better accuracies. By changing the optimizer and running the deep neural network model with different number of epochs the recognition accuracies can be varied and the model can be made to perform better [14].

Almost all the researchers working in the field of pattern recognition need to apply one or the other feature extraction technique. Kumar et al. presented a survey on indic and non-indic handwritten scripts which contained characters and numbers [3]. The survey also included numerals, handwritten alphabets and characters. The Devanagari scripts included compound conjuncts including modifiers. The methods reported in the literature were compared in the survey. Dongre et al. has compared all the existing techniques available in the field of Devanagari OCR [15]. The authors divided the complete image into multiple zones. Further calculations were done based on the percentage of black pixels present in each zone. The calculated percentage served as features for particular zone. Kimura et al. have dealt with contouring

representation of characters [16]. Orientation of contouring line segments was studied and depending upon they were classified under four sub groups: vertical, horizontal, and diagnol (45 and 135). Specifically, each countouring segment belonged to some specific group which acted as features for that segment. Prithpal et al. described the literature consisting information about OCR for Gurmukhi language recognition [105].

Alaei et al. proposed a new technique with the concept of painting. The document was decomposed into vertical stripes out of which each row was painted with the intensity pixel values. Piece wise potential separation of lines was done [17]. Arivazhagan et al. proposed a method for skewed lines which are also sometimes overlapping and worked on Arabic documents. The author traversed the lines around the handwritten documents and the complete document was divided into vertical lines. Firstly, the lines were looked upon from within 25% of the projection profile archive. The decision was taken based on the probability that was acquired by a distance metric or by evaluation of probability achieved from each Gaussian [18].

Deep learning also plays a vital role for applications like disease detection [19] and many more.

III. Research Methodology, Analysis and Results

This study used AlexNet transfer learning technique for the recognition and classification of Devanagari Characters. The experimentation was performed on 85 collected Devanagari manuscript pages. The manuscripts give the detailed death data of the persons. The collection of manuscripts were done using camera and scanner. Once the complete data was collected, pre-processing of the collected data was done, which resulted in the removal of extra add-on information like date and additional white spaces. But these pre-processed manuscripts still consisted of skewed and degraded documents. In order to remove this skewness the complete pages were first segmented into lines, words followed by characters. In the subsequent stages, the individual characters were employed for recognition purposes. The character labelling was done corresponding to the Devanagari character set. After labeling process was completed, different data augmentation techniques were applied like rotational, flip flop etc which resulted in the production of 7356 character images used for testing and training. The model was trained using AlexNet transfer learning model. In this procedure, features were extracted and employed for classification purposes. A comparison was made between the output of the transfer learning model based on AlexNet and a conventional CNN model. LabelMe data segmentation tool was used for segmenting the data into characters. LabelMe is a graphical image annotation tool which supports labeling for line, point, polygon and rectangle. It generates a single json file per image. The json file consists of list of boxes labeled by bottom-right and upper-left coordinates. Fig 4 shows the json file generated per page.

As a result of segmentation, the prepared dataset consisted of 7356 characters, set for the experimentation and labelled

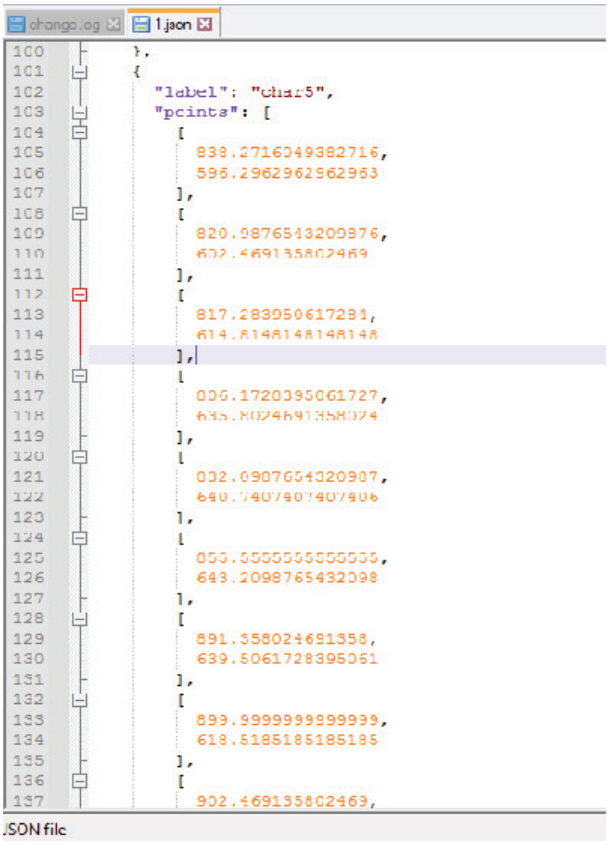

Fig. 4. JSON file generated using LabelMe

into 38 classes which comprised of some basic and conjunct characters from Devanagari script. Test and train folders consisting of 80% and 20% of data were further splitted for the experiment's result analysis, the impact of the experiment's epoch count and the size of both the testing and training data sets were taken into account. Using accuracy and loss parameters, the accuracy of the model was evaluated. Random images from dataset with labels is shown in Fig 5. The

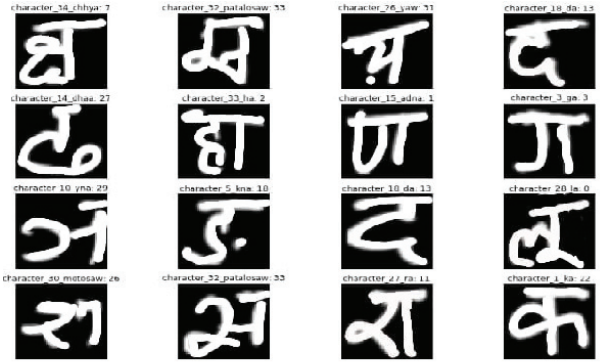

Fig. 5. Random Images With Labels

model was tested with both test and training data sizes of 90:10 (P1), 80:20 (P2), and 70:30 (P3) for each of the three

partitioning techniques. In order to evaluate the model with 10, 20, and 30 epochs, all three partitions were used. According to the P1, P2, and P3 partitions of the data set, the average accuracy values 94.96, 93.21, and 91.78 percent respectively. The combined average accuracy is 93.32 percent when all partitions and 15, 25, and 30 execution epochs are taken into account. The loss graph for validation and testing is shown in Fig 6. The combine analysis of accuracy is shown in Fig 7.

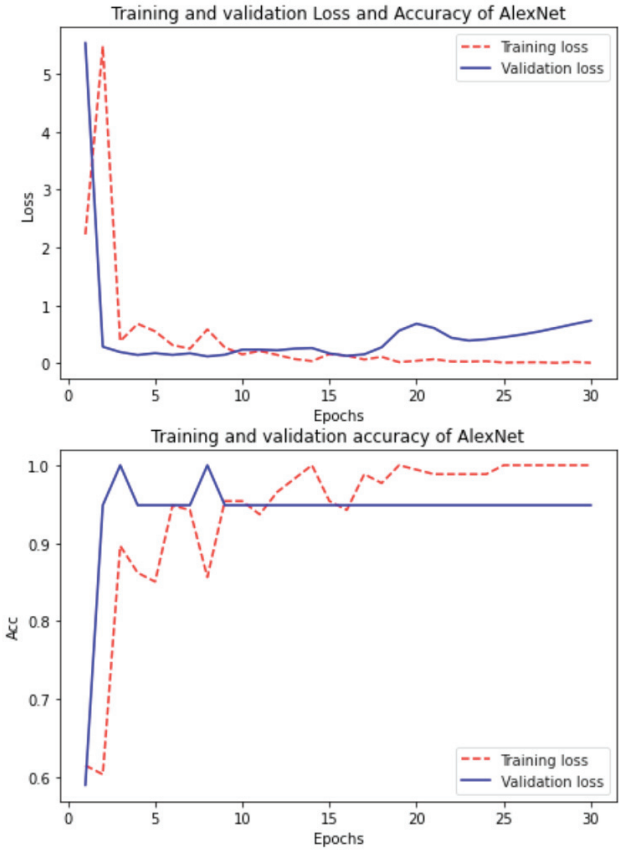

Fig. 6. Loss and accuracy analysis of AlexNet fro training and validation process for thirty Epochs

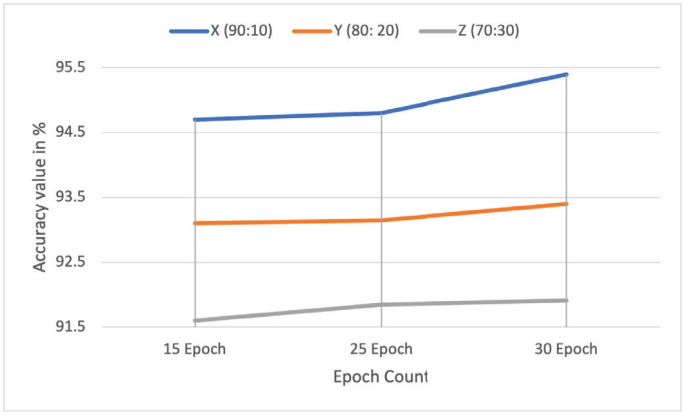

Fig. 7. Accuracy analysis

For 30 epochs with an 80:20 partition scheme, the suggested model's recall and precision values are 94.87% and 94.98% respectively. FOr the same dataset, the CNN model obtained 90.14 percent efficiency with 92.8 percent recall value and 92.19 percent precision value on P2 partition strategy. Fig 8 shows the comparison of AlexNet model with CNN model on same dataset.

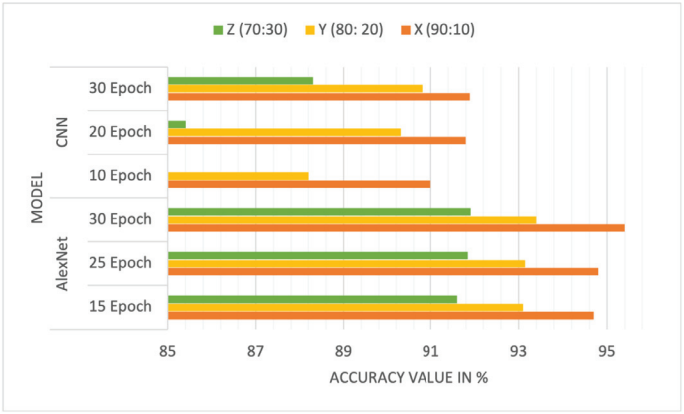

Fig. 8. Comparison based on accuracy of AlexNet and CNN based model for manuscript recognition.

IV. Conclusion

The current days offer a wide use of OCRs for easability in reading the manuscripts. But as per the literature review very less OCR systems exist for converting the regional languages into human readable forms. Furthermore, for Devanagari manuscript recognition very less amount of work has been carried out for increasing the efficacy of the system. This research paper presents one of the experiments for Devanagari manuscript recognition using alexNet. The performed experiment was able to obtain the maximum accuracy of 95.44% when experimentation was performed on 90-10 partitioning strategy. The experiments were performed taking into consideration only a few modifiers and conjuncts which could further be extended for all modifiers and conjuncts.

References

[1] N. Babu and A. Soumya, "Character recognition in historical handwritten documents–a survey," in *2019 international conference on communication and signal processing (ICCSP)*. IEEE, 2019, pp. 0299–0304.

[2] A. Moudgil, S. Singh, and B. Sareen, "Cnn-lstm based approach for recognition of devanagari manuscripts," in *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*. IEEE, 2022, pp. 1–5.

[3] M. Kumar, M. K. Jindal, R. K. Sharma, and S. R. Jindal, "Character and numeral recognition for non-indic and indic scripts: a survey," *Artificial Intelligence Review*, vol. 52, no. 4, pp. 2235–2261, 2019.

[4] R. Jayadevan, S. R. Kolhe, P. M. Patil, and U. Pal, "Offline recognition of devanagari script: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 6, pp. 782–796, 2011.

[5] A. Singh, K. Bacchuwar, and A. Bhasin, "A survey of ocr applications," *International Journal of Machine Learning and Computing*, vol. 2, no. 3, p. 314, 2012.

[6] V. Kumar and P. K. Sengar, "Segmentation of printed text in devanagari script and gurmukhi script," *International Journal of Computer Applications*, vol. 3, no. 8, pp. 30–33, 2010.

[7] E.-S. Lincke, K. Bulert, and M. Büchler, "Optical character recognition for coptic fonts: A multi-source approach for scholarly editions," in *Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage*, 2019, pp. 87–91.
[8] S. Palakollu, R. Dhir, and R. Rani, "Handwritten hindi text segmentation techniques for lines and characters," in *Proceedings of the world congress on engineering and computer science*, vol. 1, 2012, pp. 24–26.
[9] S. Singh, P. K. Sarangi, C. Singla, and A. K. Sahoo, "Odia character recognition system: A study on feature extraction and classification techniques," *Materials Today: Proceedings*, vol. 34, pp. 742–747, 2021.
[10] I. Kansal, R. Popli, and C. Singla, "Comparative analysis of various machine and deep learning models for face mask detection using digital images," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*. IEEE, 2021, pp. 1–5.
[11] Ø. D. Trier, A. K. Jain, and T. Taxt, "Feature extraction methods for character recognition-a survey," *Pattern recognition*, vol. 29, no. 4, pp. 641–662, 1996.
[12] J.-B. Yang, K.-Q. Shen, C.-J. Ong, and X.-P. Li, "Feature selection for mlp neural network: The use of random permutation of probabilistic outputs," *IEEE Transactions on Neural Networks*, vol. 20, no. 12, pp. 1911–1922, 2009.
[13] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 609–616.
[14] D. Khanduja, N. Nain, and S. Panwar, "A hybrid feature extraction algorithm for devanagari script," *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 15, no. 1, pp. 1–10, 2015.
[15] V. J. Dongre and V. H. Mankar, "A review of research on devnagari character recognition," *arXiv preprint arXiv:1101.2491*, 2011.
[16] F. Kimura and M. Shridhar, "Handwritten numerical recognition based on multiple algorithms," *Pattern recognition*, vol. 24, no. 10, pp. 969–983, 1991.
[17] A. Alaei, P. Nagabhushan, and U. Pal, "Piece-wise painting technique for line segmentation of unconstrained handwritten text: a specific study with persian text documents," *Pattern Analysis and Applications*, vol. 14, no. 4, pp. 381–394, 2011.
[18] M. Arivazhagan, H. Srinivasan, and S. Srihari, "A statistical approach to line segmentation in handwritten documents," in *Document recognition and retrieval XIV*, vol. 6500. SPIE, 2007, pp. 245–255.
[19] P. Nagu, A. Parashar, T. Behl, and V. Mehta, "Cns implications of covid-19: a comprehensive review," *Reviews in the Neurosciences*, vol. 32, no. 2, pp. 219–234, 2021.